

# **Database Searching Tips**

May 24, 2000

Barb O'Bryen, STIC

- I. Inventor Searches
- II. Database Selection
- III. Database Clusters
- IV. Choosing Terms to Use in Your Search
- V. Searching for Chemicals
- VI. Narrowing Your Search
- VII. Indexing
- VIII. Removing Duplicates
- IX. Display and Print Formats
- X. A Few "Don'ts"
- XI. Where to Turn for Help

## Inventor Searches

### *STN*

s ((zurkowitz b?) or (zurkowitz, b?))/au

### *Dialog*

s au=((zurkowitz, b?) or (zurkowitz b?))

Spanish, Dutch, and Swedish surnames are not treated in a consistent manner in all databases; multiple variations should be used to ensure complete retrieval. For example, to find works written by Carmelita Rodriguez Otero, search these variants:

### *STN*

s ((rodriguez otero c?) or (rodriguez otero, c?) or (otero c?) or (otero, c?))/au

### *Dialog*

s au=((rodriguez otero, c?) or (rodriguez otero c?) or (otero, c?) or (otero c?))

Irish and Scottish surnames may also be problematic.

### *STN*

s ((obryen b?) or (obryen, b?) or (o bryen b?) or (o bryen b?))/au  
s ((mackenzie i?) or (mackenzie, i?) or (mac kenzie i?) or (mac kenzie, i?))/au

### *Dialog*

s au=((obryen, b?) or (obryen b?) or (o bryen, b?) or (o bryen b?))  
s au=((mackenzie, i?) or (mackenzie i?) or (mac kenzie i?) or (mac kenzie, i?))

## Database Suggestions

### Basic General Biotechnology Search

Medline  
Biosis  
CAPlus  
Embase  
Derwent World Patents Index

### Expanded General Biotechnology Search

*consider adding some of these databases to the list above:*

Chemical Engineering and Biotechnology Abstracts  
Conference Papers Index  
Current Biotechnology Abstracts  
Derwent Biotechnology Abstracts  
Dissertation Abstracts  
Inside Conferences  
JAPIO  
JICST-E  
Life Sciences Collection  
NTIS  
Pascal  
World Translations Index

### Agriculture/Veterinary Search

Agricola  
AGRIS International  
CRIS/USDA  
CAB Abstracts  
CAPlus  
Biosis  
Pascal  
Medline

### Food Sciences Search

Agricola  
CRIS/USDA  
Food Science and Technology Abstracts  
Foods Adlibra  
CAB Abstracts  
Biosis

### Pollution Bioremediation Search

Biosis  
CAB Abstracts  
CAPlus  
Enviroline  
Environmental Bibliography  
Pollution Abstracts  
NTIS  
GeoArchive  
Aquatic Science and Fisheries Abstracts  
Oceanic Abstracts  
Waternet  
Water Resources Abstracts

### Alternative/Natural Medicine Search

Medline  
Natural Products Alert  
Embase  
CAPlus  
Biosis  
MANTIS  
Cancerlit  
AIDSLINE  
Derwent World Patents Index  
JAPIO  
Chinese Patent Abstracts

### AIDS/HIV Search

search AIDSLINE in place of Medline

### Cancer Search

search Cancerlit in place of Medline

## Selected Databases Available on STN and Dialog

Database Name	Dialog file number	STN file name
Agricola	10	AGRICOLA
AIDSline	157	AIDSLINE
Analytical Abstracts	305	ANABSTR
Aquatic Sciences and Fisheries Abstracts	44	AQUASCI
Biosis Previews	5	BIOSIS
CAB Abstracts	50	CABA
Cancerlit	159	CANCERLIT
Chemical Engineering and Biotechnology	315	CEABA
Conference Papers Index	77	CONFSCI
Derwent Biotechnology Abstracts	357	BIOTECHDS
Embase	73	EMBASE
Energy Science and Technology (aka Energy)	103	ENERGY
Food Science and Technology Abstracts	51	FSTA
GeoRef	89	GEOREF
International Pharmaceutical Abstracts	74	IPA
JAPIO	347	JAPIO
JICST-E	94	JICST-E
Life Sciences Collection	76	LIFESCI
Medline	155	MEDLINE
NTIS	6	NTIS
Oceanic Abstracts (aka Ocean)	28	OCEAN
Pollution Abstracts	41	POLLUAB
SciSearch	434	SCISEARCH

### Selected Databases Exclusive to STN

Database Name	STN file name
Chemical Abstracts (with abs)	CAPLUS
Chemical and Engineering News Online	CEN
Chemical Journals of Elsevier Science Publishers (fulltext)	CJELSEVIER
Chemical Journals of John Wiley & Sons, Inc (fulltext)	CJWILEY
Chemical Journals of the American Chemical Society	CJACS
Chemical Journals of the AOAC International (fulltext)	CJAOAC
Chemical Journals of the Royal Society of Chemistry (fulltext)	CJRSC
COMLINE Daily News	JPNEWS
Cosmetic and Perfume Science Technology	KOSMET
Health and Safety Science Abstracts	HEALSAFE
Natural Products Alert	NAPRALERT
Registry (with str searching)	REGISTRY
ULIDAT	ULIDAT

### Selected Databases Exclusive to Dialog

Database Name	Dialog file
AGRIS	203
Biological and Agricultural Index	143
CRIS/USDA	60
Current Biotechnology Abstracts	358
Dissertation Abstracts Online	35
Environmental Bibliography	68
Foodline: Food Science & Technology	53
Foods Adlibra	79
GeoArchive	58
Inside Conferences	65
Water Resources Abstracts	117
Waternet	245
World Translations Index	295

# BIOTECHNOLOGY & CHEMICAL NEWS

## DIRECTORY ASSISTANCE FOR ONLINE SEARCHING

When searching for an unknown phone number, you instinctively call directory assistance. What can you do when looking for appropriate databases in which to conduct your STN or Dialog online search? Fortunately, both vendors have directory assistance-like tools designed for this purpose. In STN, the tool is called STNINDEX, and in Dialog it is aptly named File 411. The example below illustrates just how these tools can save you time and money.

### STNINDEX

STN has created clusters of files with related properties, a list of which can be found at the back of STN database catalog. You can use these clusters or create your own using the "set cluster" command. There is no limit to the number of files in a cluster. The cost of this index searching is a modest \$45 per hour. Using the index, you can identify the best databases for your desired search, or you can formulate a whole series of query statements, each with its unique L#, and then you can combine the L# using standard Boolean logic.

The search outlined below aimed to identify the top ten bioscience databases that contained the term EBOLA.

=> index bioscience

INDEX 'AIDSLINE, ANABSTR, AQUASCI, BIOBUSINESS, BIOSIS, BIOTECHABS, BIOTECHDS, CABA, CANCERLIT, CAPLUS, CEABA, CEN, CIN, CJACS, CJELSEVIER, CONFSCI, DDFB, DDFU, DISSABS, DRUGB, DRUGLAUNCH, DRUGNL, DRUGU, EMBAL, EMBASE, FSTA, GENBANK, HEALSAFE, IFIPAT, ...' ENTERED AT 11:20:52 ON 26 DEC 95

42 FILES IN THE FILE LIST IN STNINDEX; Enter SET DETAIL ON to see search term postings or to view

=> s ebola

FILE 'AIDSLINE'

4 EBOLA

:

FILE 'USPATFULL'

17 EBOLA

L1 QUE EBOLA

=> d rank

F1	210	MEDLINE
F2	173	EMBASE
F3	163	SCISEARCH
F4	156	BIOSIS
F5	75	LIFESCI
F6	44	NTIS
F7	32	CABA
F8	27	CAPLUS
F9	26	PROMT
F10	25	TOXLINE

:

F29 1 FSTA

=> d cost

COST IN U.S. DOLLARS  
FULL ESTIMATED COST

0.45 0.75

The L number thus created can now be easily searched in all or any of the ranked files. You may start up all the ranked files which contained your hit term by simply typing:

=> file hits

alternatively, desired files can be selected by using the F# displayed in the ranking:

=> file f1-4 f10.

Once the selected files start up, simply type:

=> search L1

to retrieve all the records.

Remember to remove duplicate records by using the "DUP REM L#" command



## Creating Personalized Database Clusters

### *Dialog*

? b 155

File 155:MEDLINE(R) 1966-1999/May W2

(c) format only 1999 Dialog Corporation

\*File 155: reloaded, note accession numbers changed.

Set Items Description

--- -----

? recall profile

Line Command

-----

1. SET HI\*\*\*\*
2. SET KWIC 50
3. SET BANNERS SHORT
4. SET COST SHORT

? edit profile

Editor entered

Name: PROFILE

Total lines: 4

Line increment: 10

Last line: 40

EDIT:

? i

INPUT: 50

? set alias vetrnry 60,155,10,203,50,5

INPUT: 60

? <enter blank line>



Returning to EDIT mode  
EDIT:

? save profile

-->Replace "PROFILE"? ( Y = Yes N = No )

? yes

PROFILE stored  
Exit from editor

? logoff

Welcome to DIALOG

Dialog level 99.02.26D

Last logoff: 22mar99 17:57:46

Logon file405 22mar99 17:58:10

dialog

HILIGHT set on as '\*\*\*\*'

KWIC is set to 50.

Banner display set SHORT.

COST = SHORT.

VETRNRY is set ON as an alias for 60,155,10,203,50,5.

? b vetrnry

SYSTEM:OS - DIALOG OneSearch

File 60:CRIS/USDA (c) format only 1998 The Dialog Corporaion pl

\*File 60: This file will no longer be updated.

File 155:MEDLINE(R) (c) format only 1999 Dialog Corporation

\*File 155: reloaded, note accession numbers changed.

File 10:AGRICOLA (c) format only 1999 The Dialog Corporation

File 203:AGRIS Dist by NAL, Intl Copr. All rights reserved

File 50:CAB Abstracts (c) 1999 CAB International

File 5:BIOSIS PREVIEWS(R) (c) 1999 BIOSIS

(To see coverage dates, enter SHOW FILES.)

Set Items Description

--- -----

## Search and Display Costs

### Dialog Biotechnology and Chemical Files

File Name	DIALOG file number	RAW COSTS*		Approx.
		DialUnit	Display All	Overall Cost
CRIS/USDA	60	\$1.25	\$0.70	\$19.50
AIDSLINE®	157	\$2.25	\$0.25	\$25.00
CANCERLIT®	159	\$2.75	\$0.25	\$30.00
MEDLINE®	155	\$3.00	\$0.20	\$32.00
Biological & Agricultural Index	143	\$2.25	\$1.05	\$33.00
Environmental Bibliography	68	\$2.25	\$1.05	\$33.00
AGRICOLA	10	\$2.10	\$1.20	\$33.00
AGRIS International	203	\$2.25	\$1.25	\$35.00
CAB ABSTRACTS	50	\$2.75	\$1.50	\$42.50
JICST-Eplus- Japanese Science & Technology	94	\$3.25	\$1.15	\$44.00
Inside Conferences	65	\$3.50	\$1.00	\$45.00
Conference Papers Index	77	\$2.75	\$1.80	\$45.50
PASCAL	144	\$3.25	\$1.45	\$47.00
Foodline®: Food Science & Technology	53	\$4.00	\$1.60	\$56.00
Food Science and Technology Abstracts	51	\$4.25	\$1.60	\$58.50
Dissertation Abstracts Online	35	\$4.00	\$2.10	\$61.00
Pollution Abstracts	41	\$4.50	\$1.65	\$61.50
Energy Science and Technology	103	\$4.75	\$1.45	\$62.00
Life Sciences Collection	76	\$4.75	\$1.65	\$64.00
Current Biotechnology Abstracts	358	\$3.50	\$2.90	\$64.00
Aquatic Sciences and Fisheries Abstracts	44	\$5.00	\$1.65	\$66.50
GeoRef	89	\$5.75	\$1.00	\$67.50
BIOSIS Previews®	5	\$5.25	\$1.55	\$68.00
Oceanic Abstracts	28	\$5.50	\$1.65	\$71.50
Enviroline®	40	\$6.50	\$1.10	\$76.00
International Pharmaceutical Abstracts	74	\$5.75	\$1.90	\$76.50
USPATFULL	652-654	\$5.50	\$3.00	\$85.00
Chemical Engineering and Biotechnology Abstracts	315	\$6.50	\$2.40	\$89.00
EMBASE®	73	\$7.75	\$1.90	\$96.50
EMBASE® Alert	172	\$7.75	\$2.25	\$100.00
Analytical Abstracts	305	\$7.25	\$3.30	\$105.50
European Patents Fulltext	348	\$4.25	\$6.60	\$108.50
Derwent Biotechnology Abstracts	357	\$8.78	\$2.08	\$108.60
JAPIO	347	\$10.25	\$1.40	\$116.50
SciSearch®: A Cited Reference Science Database	434	\$11.85	\$3.00	\$148.50
Derwent World Patents Index	351	\$16.25	\$3.55	\$198.00
IMSworld Product Launches	446	\$14.25	\$5.60	\$198.50
Pharmaceutical and Healthcare Industries News Archive	129	\$23.00	\$4.25	\$272.50
Pharmaceutical and Healthcare Industries News Current	130	\$23.00	\$8.50	\$315.00
IMSworld R&D Focus Drug News	445	\$14.25	\$36.90	\$511.50

\*Cost information current as of 1/11/99

STN

=> fil home

FILE 'HOME' ENTERED AT 17:34:22 ON 22 MAR 1999

=> set cluster

ENTER CLUSTER NAME OR (?):.biotech

ENTER LIST OF FILE NAMES OR (?):medline,biosis,biotechds,caplus,embase

MORE FILES, (NONE) OR ?:

*choice that appears in parenthesis can be selected by typing a period*

CLUSTER '.BIOTECH' DEFINED AS 'MEDLINE, BIOSIS, BIOTECHDS, CAPLUS, EMBASE'

SET COMMAND COMPLETED

*Cluster names must:*

*begin with a period ( . )*

*have 2-16 characters*

*contain only letters, numbers, and periods*

=> file .biotech

FILE 'MEDLINE' ENTERED AT 17:36:19 ON 22 MAR 1999

FILE 'BIOSIS' ENTERED AT 17:36:19 ON 22 MAR 1999  
COPYRIGHT (C) 1999 BIOSIS(R)

FILE 'BIOTECHDS' ENTERED AT 17:36:19 ON 22 MAR 1999  
COPYRIGHT (C) 1999 DERWENT INFORMATION LTD

FILE 'CAPLUS' ENTERED AT 17:36:19 ON 22 MAR 1999  
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
COPYRIGHT (C) 1999 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'EMBASE' ENTERED AT 17:36:19 ON 22 MAR 1999  
COPYRIGHT (C) 1999 Elsevier Science B.V. All rights reserved.

=>

=> display cluster

ENTER (ENTIRE), OR A LIST OF CLUSTER NAMES:.

CLUSTER NAME    COMMENT

-----  
.BARBS

.BIOTECH

AEROTECH      Aerospace and Related Technology Cluster

AGRICULTURE    Agriculture Cluster

.  
.  
.

SUPPLIERS      Product Directories and Suppliers Cluster

TOXICOLOGY    Toxicological Information Cluster

TO SEE THE LIST OF FILES THAT COMPRISE A CLUSTER, ENTER 'DISPLAY CLUSTER' FOLLOWED BY THE CLUSTER NAME. ENTER 'HELP CLUSTER NAMES' TO SEE A LIST OF ALL SYSTEM DEFINED CLUSTERS AND THEIR DEFINITIONS.

=> display cluster .biotech

CLUSTER NAME    CLUSTER DEFINITION

-----  
.BIOTECH      MEDLINE BIOSIS BIOTECHDS CAPLUS EMBASE

*Online guidance for creating database clusters on STN can be obtained by typing help set cluster at the system prompt.*

# **Search and Display Costs** **STN Biotechnology and Chemical Files**

File Name	STN File Name	RAW COSTS*			Approx. Overall Cost
		Connect Hour	Search Term	Display All	
MEDLINE	MEDLINE	\$30.00		\$0.17	\$ 16.70
AIDS Literature Online	AIDSLINE	\$30.00		\$0.20	\$ 17.00
Cancer Literature Online	CANCERLIT	\$30.00		\$0.20	\$ 17.00
Agriculture Online Access	AGRICOLA	\$20.00		\$1.15	\$ 21.50
CAB ABSTRACTS	CABA	\$30.00		\$1.45	\$ 29.50
JICST-EPlus	JICST-EPlus	\$45.00		\$1.10	\$ 33.50
Geological Reference	GeoRef	\$50.00		\$0.95	\$ 34.50
International Pharmaceutical Abstracts	IPA	\$45.00		\$1.50	\$ 37.50
DOE ENERGY	ENERGY	\$52.00		\$1.48	\$ 40.80
Food Science and Technology Abstracts	FSTA	\$52.00		\$1.54	\$ 41.40
Chemical Abstracts Plus (Search Term Rate)	ZCAplus	\$0.00	\$1.73	\$2.00	\$ 45.95
Biosis Previews®/RN	BIOSIS	\$60.00		\$1.60	\$ 46.00
Derwent Biotechnology Abstracts	BIOTECHDS	\$83.00		\$1.23	\$ 53.80
Chemical Abstracts Plus (Search Term + Hourly Rate)	CAPLUS	\$30.00	\$1.33	\$2.11	\$ 56.05
Conference Papers Index	CONFSCI	\$80.00		\$1.65	\$ 56.50
Life Sciences Collection	LIFESCI	\$80.00		\$1.65	\$ 56.50
Excerpta Medica	EMBASE	\$83.00		\$1.65	\$ 58.00
Predicasts Overview of Markets and Technology	PROMT	\$60.00		\$2.85	\$ 58.50
Japanese Patent Information Organization	JAPIO	\$88.00		\$1.48	\$ 58.80
Aquatic Sciences and Fisheries Abstracts	AQUASCI	\$90.00		\$1.80	\$ 63.00
Oceanic Abstracts	OCEAN	\$90.00		\$1.80	\$ 63.00
Pollution Abstracts	POLLUAB	\$90.00		\$1.80	\$ 63.00
EMBASE ALERT	EMBAL	\$90.00		\$2.20	\$ 67.00
REGISTRY (Search Term Rate)	ZREGISTRY	\$0.00	\$4.00	\$1.35	\$ 73.50
Chemical Engineering and Biotechnology Abstracts	CEABA	\$112.00		\$2.24	\$ 78.40
SciSearch®	SCISEARCH	\$105.00		\$2.85	\$ 81.00
REGISTRY (Search Term + Hourly Rate)	REGISTRY	\$30.00	\$3.55	\$1.35	\$ 81.75
Natural Products Alert	NAPRALERT	\$50.00		\$6.45	\$ 89.50
USPATFULL	USPATFULL	\$105.00		\$3.80	\$ 90.50
IFI Patent Database	IFIPAT	\$130.00		\$2.85	\$ 93.50
IMSworld R&D Focus Drug News	DRUGNL	\$145.00		\$2.80	\$ 100.50
Chemical Abstracts Plus (Hourly Rate)	HCAplus	\$171.00		\$2.11	\$ 106.60
Analytical Abstracts	ANABSTR	\$158.00		\$3.38	\$ 112.80
IMSworld Drug Launch	DRUGLAUNCH	\$145.00		\$5.40	\$ 126.50
Pharmaceutical and Healthcare Industries News Archive	PHIN	\$210.00		\$2.20	\$ 127.00
Derwent World Patent Index	WPIDS	\$266.00		\$2.44	\$ 157.40
Pharmaceutical and Healthcare Industries News Current	PHIC	\$210.00		\$7.35	\$ 178.50

\* Cost information current as of 1/11/99

\*\*Additional charge for records with sequences

## Construct Your Search Strategy to Allow for Vocabulary Variability

Truncate words to save typing

*anor?* will retrieve anorectic, anorexia, anorexiant, anorexic, anorexigenic

*inhibit?* will retrieve inhibit, inhibitor, inhibitors, inhibition, inhibited, inhibiting

*the CAPLUS file on STN also offers left truncation - ?neoplas? will retrieve antineoplastic, antineoplastics, preneoplasm, preneoplastic, neoplastic, neoplasm, neoplasms, neoplastic*

Include synonyms, near-synonyms, and common abbreviations

antibod? or immunoglobulin?

newborn? or infant? or neonat?

cell(w)death or apopto?

detect? or determin? or anal? or measure? or assay? or quantit?

inhibit? or prevent? or disrupt? or block? or modul? or reduc? or decreas? or suppress?

heat(3a)(resist? or stab?) or thermostab? *STN*

heat(3n)(resist? or stab?) or thermostab? *Dialog*

ELISA or EIA or immunoassay? or immuno(W)assay?

(modulat? or time# or delay? or sustain? or control?)(3a)(deliver? or releas? or action or acting) *STN*

(modulat? or time? ? or delay? or sustain? or control?)(3n)(deliver? or releas? or action or acting) *Dialog*

arteriosclero? or artherosclero?

cancer? or neoplas? or tumor? or tumour? or malignan? [also consider using names of specific types of cancer, such as leukemia, sarcoma, myeloma, etc.]

Antonyms may also turn up useful references

anore? or hypoglycem? or ((weight or wt)(a)(los### or body or gain)) or appetite or obese or obesity *STN*

anore? or hypoglycem? or ((weight or wt)(n)(los??? or body or gain)) or appetite or obese or obesity *Dialog*

sunscreen? or sunblock? or suntan? or sunburn? or UV# or ultraviolet? or ultra(W)violet *STN*

sunscreen? or sunblock? or suntan? or sunburn? or UV? ? or ultraviolet? or ultra(W)violet *Dialog*

Use common names and scientific names for plants and animals

tick# or amblyomma or argas or boophilus or dermacentor or haemaphysalis or ornithodoros or  
rhipicephalus or ixod? [family] or argasid? [family] **STN**

tick? ? or amblyomma or argas or boophilus or dermacentor or haemaphysalis or ornithodoros or  
rhipicephalus or ixod? [family] or argasid? [family] **Dialog**

*cast an even wider net by moving into broader biological classifications*

acarid# [order] or arachnid? [class] or arthropod? [phylum] **STN**

acarid? ? [order] or arachnid? [class] or arthropod? [phylum] **Dialog**

Remember variable abbreviations

(molec? or mol)(W)(weight or wt) or MW

kda or kilodalton? or dalton? or kd

*STN has a feature which will automatically search all abbreviations of a word in the basic indexes of  
English-language databases.*

*In File Home, enter the command => set abb on perm*

=> fil hom

FILE 'HOME' ENTERED AT 15:44:58 ON 06 JUL 1998

=> set abb on perm

SET COMMAND COMPLETED

=> fil capl

FILE 'CAPLUS' ENTERED AT 15:45:08 ON 06 JUL 1998

.  
.  
.

=> s concentrat?

160736 CONCENTRAT?

53974 CONC

121622 CONCD

10336 CONCG

1117639 CONCN

L5 1315811 CONCENTRAT?

(CONCENTRAT? OR CONC OR CONCD OR CONCG OR CONCN)

## Consider British spellings

aluminium rather than aluminum

sulphur, sulphuric, sulphate rather than sulfur, sulfuric, sulfate

'c' may replace 'k':

mollusc rather than mollusk

leucocyte rather than leukocyte

*use internal masking to search either variation of leukocyte:*

leu!ocyte *STN*

leu?ocyte *Dialog*

's' replaces 'z' in many words:

analyse

hybridise

immunise

nebulise

neutralise

polymerise

pulverise

stabilise

sterilise

*truncate terms after the 'y' or 'i' preceeding the 's/z'*

'or' within a word becomes 'our':

colour

tumour

vapour

'e' becomes 'ae' or 'oe':

anaesthetise

chimaeric

haematin

haemorrhage

leukaemia

amoeba

foetal

oestrogen



## Searching for Chemicals Using Registry Numbers

Some files useful for biotechnology searching contain Registry Numbers, which can make searching for chemicals easier. A partial list of files with Registry Numbers includes:

### On STN

AIDSline  
Analytical Abstracts  
Biosis  
Cancerlit  
CAPlus  
Embase  
Medline  
Natural Products Alert  
US Patents Fulltext

### On Dialog

AIDSline  
Analytical Abstracts  
Cancerlit  
Embase  
International Pharmaceutical Abstracts  
Medline  
(Biosis is available on Dialog, but does not contain Registry Numbers)

The two best sources for finding the Registry Number of a compound are:

paper - *The Merck Index: An Encyclopedia of Chemicals and Drugs*  
RS356.M524 located in the reference section of the library

online - the Registry file on STN

=> fil reg

FILE 'REGISTRY' ENTERED AT 14:57:26 ON 15 JUL 1998

=> e edta/cn      *expand on the chemical name*

E1	1	EDT 69N-S/CN
E2	1	EDT-TTF (SC18) 2/CN
E3	1 -->	EDTA/CN
E4	1	EDTA (3-) /CN
E5	1	EDTA (CHELATING AGENT) /CN
E6	1	EDTA 10/CN
E7	1	EDTA 10P/CN
E8	1	EDTA AMMONIUM SALT/CN
E9	1	EDTA AMYLAMINE SALT/CN
E10	1	EDTA BUTYLAMINE SALT/CN
E11	1	EDTA CALCIUM DISODIUM SALT/CN
E12	1	EDTA CALCIUM SALT/CN

=> s e3

L1      1 EDTA/CN      *if found, search on the appropriate "e-number"*

=> d rn cn     *display the Registry Number and the chemical names*

L1    ANSWER 1 OF 1    REGISTRY    COPYRIGHT 1998 ACS  
RN    60-00-4    REGISTRY  
CN    Glycine, N,N'-1,2-ethanediylbis[N-(carboxymethyl)- (9CI)  
      (CA INDEX NAME)  
OTHER CA INDEX NAMES:  
CN    Acetic acid, (ethylenedinitrilo)tetra- (8CI)  
OTHER NAMES:  
CN    3,6-Diazaoctanedioic acid, 3,6-bis(carboxymethyl)-  
CN    Acetic acid, 2,2',2'',2'''-(1,2 ethanediyl dinitrilo)  
      tetrakis-  
      ~~  
CN    Edetic acid  
CN    **EDTA**  
CN    EDTA (chelating agent)  
CN    Endrate  
CN    Ethylenediamine-N,N,N',N'-tetraacetic acid  
CN    Ethylenediaminetetraacetic acid  
CN    Ethylenedinitrilotetraacetic acid  
      ~~  
CN    YD 30

Use both the Registry Number and one or more of the chemical names in your search strategy.

### Dialog

Search the Registry Number with search prefix and chemical names.

?s rn=60-00-4  
?s edta  
?s ethylenediaminetetraacetic  
?s ethylenediamine(S)tetraacetic

### STN

Search the L-number from the Registry file which represents your compound, plus chemical names.

=> s L1        *L1 was the Registry file answer set from the example on the previous page*  
=> s edta  
=> s ethylenediaminetetraacetic  
=> s ethylenediamine(s)tetraacetic

## Searching for Chemicals in Files Lacking Registry Numbers

In these files, you must rely entirely on text searching of the chemical name. Include as many synonyms as you deem practical. Although you will not be able to use Registry Numbers in these files, a search of the compound in the Registry file can still be worthwhile as a way of finding those synonyms.

Be aware that Derwent World Patents often breaks up chemical names into smaller parts and make allowances for this in your search strategy. For example, the chemical ethylenediaminetetraacetic acid should be searched as:

(ethylenediamine or (ethylene(w)(diamine or (di(w)amine))))(S)(tetraacetic or (tetra(w)acetic))  
or ethylenediaminetetraacetic or edta

## Ways to Narrow a Search

- 1) Scan some of your answers to see if one particular term is causing most of the false drops; eliminate that term from your strategy.
- 2) Use proximity operators or "S" instead of "and" to combine terms.

"S" requires terms to be in the same field, but does not specify which field.

=> s carrot?(S)(soft(W)rot#####) *STN*

? s carrot?(S)(soft(W)rot????) *Dialog*

Proximity operators require terms to appear within a specified number of words of each other.

=> s embryo?(5a)cultur? *STN*

? s embryo?(5n)cultur? *Dialog*

- 3) Limit retrieval to English language articles only.

=> s L10 and english/LA *STN*

? s S10/ENG *Dialog*

- 4) Limit by publication date.

nothing more recent than a given date:

=> s L5 not py>1995 *STN*

? s S5 not py>1995 *Dialog*

Published within a range of dates:

=> s L5 and 1993-1995/py

? s S5 and py=1993:1995

## Using Controlled Vocabulary Adds Precision to Your Search

Print thesauri are available in the library and in room 12C14 for:

### Biosis

*Biosis Previews Search Guide* QH303.B567

### Chemical Abstracts

*Chemical Abstracts Index Guide* QD1.A51

### Embase

*EMTREE Thesaurus - Alphabetical Index* Z699.5.M39 E522  
*EMTREE Thesaurus - Tree Structure* Z699.5.M39 E522

### Medline (MeSH)

*Medical Subject Headings - Annotated Alphabetic List* R100.I53  
*Permuted Medical Subject Headings* R100.I55  
*Medical Subject Headings - Tree Structures* R100.I58

## Syntax for Including Controlled Vocabulary in a Search

### Biosis

s cc=12512

to "major" a concept code, search the S-number/MAJ: ?s S3/MAJ *Dialog*

s 12512/cc

to "major" a concept code, precede the code with an asterisk: =>s \*12512/cc *STN*

### CAPlus

s neoplasm inhibitors/ct

s "fungicides and fungistats"/ct *if the controlled term includes the word 'and', the entire term must be enclosed in quotation marks*

s therapeutics/ct (L) geno *controlled terms can be combined with free text*

### Embase

s systemic lupus erythematosus/de

s autoimmunity! *search term plus all terms under it in the tree (more specific terms)*

to "major" a concept code, search the S-number/MAJ: ?s S3/MAJ *Dialog*

s systemic lupus erythematosus/ct

s autoimmunity+nt/ct *search term plus all terms under it in the tree (more specific terms)*

to "major" a concept code, search the L-number/MAJ: =>s L3/MAJ *STN*

### Medline

s amyotrophic lateral sclerosis/de

s neuromuscular diseases! *explodes the term*

to "major" a concept code, search the S-number/MAJ: ?s S3/MAJ *Dialog*

s amyotrophic lateral sclerosis/ct

s neuromuscular diseases+nt/ct *explodes the term*

to "major" a concept code, search the L-number/MAJ: =>s L3/MAJ *STN*

## Removing Duplicates

Both STN and Dialog have duplicate removal functions.

### *Dialog - rd*

Duplicates can be removed from multifile search answer sets only.

? rd S14

### *STN - dup rem*

Duplicates can be removed from a single L-number in a multifile search, or from the answer sets of searches done in several individual files.

Multifile search example: => dup rem L7

To remove duplicates from answer sets of searches done individual files, include each appropriate L-number, separated by commas, in the command. L-numbers need not be in numerical order. List the L-number of the answer set from the least expensive file first.

Individual file example: => dup rem L10, L53, L20

## Display and Print Formats

Both vendors offer inexpensive or free formats for **scanning** results.

### STN

- CAPlus      => d scan L# (shows title and indexing) *free!*  
              => d scan ti L# (shows title only)      *free!*
- Biosis       => d scan L# (shows title and keywords) *free!*  
              => d ti cc L# [answer numbers] (shows title and indexing)  
                  e.g. => d ti cc L25 1-7
- Medline      => d trial L# [answer numbers] (shows title and indexing) *free!*
- Embase       e.g. => d trial L5 1-8

For most other files on STN, the free display format is => d scan

When doing a multifile search, the format which is free for some files and inexpensive for others, but which works for all files, is => d ti L# [answer numbers] e.g. => d ti L20 1-15

### Dialog

Within a file, you can determine what display formats are available by typing the command:  
? help fmt [file number] e.g. ? help fmt 5 to get a listing of display formats for Biosis (file 5).  
Unfortunately, no cost information is displayed.

The standard free or inexpensive displays in almost all Dialog files are:

- ? t S#/8/[answer numbers] e.g. ? t S23/8/1-5 (shows title and indexing)
- ? t S#/6,k/[answer numbers] e.g. ? t S18/6,k/1-5 (shows title and displays a window of  
text around each of your search terms)
- ? t S#/free/[answer numbers] e.g. ? t S18/free/1-5 (information displayed varies from file to file,  
but there is no charge for the display)

When you have completed your search and are ready to display the **full answers**, activate a transcript file, then use the following formats, which both provide the complete bibliographic citation and full abstract:

### STN

=> d bib ab L# [answer numbers] e.g. => d bib ab L54 1-25

### Dialog

? t S#/7/[answer numbers] e.g. ? t S54/7/1-25



## A Few Don'ts

- Don't search Chemical Abstracts on Dialog (CASearch; File 399)
- Don't *automatically* limit by date.
- Don't forget cost

in multifile searches, put the cheapest files at the beginning of the list

get cost information online

=> help cost (enter command in the file of interest) *STN*

? help cost (to get abbreviated cost information for all files) *Dialog*

? help rates [file number] (for detailed cost information for a specific file)

e.g. ? help rates 155

get cost information on the Web

### *Dialog*

<http://library.dialog.com/bluesheets>

*bluesheets, for detailed cost information for a specific database*

[http://products.dialog.com/products/dialog/dial\\_pricing.html](http://products.dialog.com/products/dialog/dial_pricing.html) "Database Rates"

*a PDF file with cost information for all files*

### *STN*

<http://www.cas.org/stn.html> "1999 STN (Columbus) Price List"

*a PDF file with cost information for all files*

- When submitting a search to STIC:

Don't list keywords with no indication of their relationship to each other

Don't attach claims without any explanation

## Where to Turn for Help with Database Searching

### STIC

*First floor, at the back of the library:*

John Dantzman	1E05	308-4488
Jan Delaval	1E01	308-4498
Toby Port	1E01	308-3534
Paula Sheppard	1E03	308-4499

*Twelfth floor:*

Mary Hale	12D16	308-4258
Susan Hanley	12C14	308-4053
Edward Hart	12C14	208-9203
Barb O'Bryen	12C14	308-4291
David Schreiber	12E18	308-4292
Beverly Shears	12C14	308-4994
Mona Smith	12C14	308-3278
Alex Waclawiw	12C14	308-4261 (works Mon and Thurs only)

### Vendor Help Desks

Dialog	800-234-2564	24 hours M-F; 10 AM to 6 PM Saturday	(have IDs. no hardy)
STN	800-848-6533	8 AM to 8 PM, M-F	

### Information Technology Resource Providers

Lora Burgess	3E18	308-4592 Pager: 616-1035 (enter your phone number and # after the beep)
Vivian Owens	5D14	308-3330 Pager: 512-3965

## System Command Quick Reference

	<u>STN</u>	<u>Dialog</u>
<b>System Prompt</b>	=>	?
<b>Change Files</b>	FILE file medline	B b 155
<b>Execute a Search</b>	S s banana	S s banana
<b>Answer Set</b>	L# L2	S# S2
<b>Truncation</b>		
Unlimited	? neoplas?	? neoplas?
0-1 Characters	# cat#	? ? cat? ?
1 character internal (can be combined with truncation at end of word)	! leu!ocyte	? leu?ocyte
<b>Proximity operator</b>	A antibod?(a)monoclonal heat(3a)(resist? or stab?)	N antibod?(n)monoclonal heat(3n)(resist? or stab?)
<b>Look at Results</b>	D D L3 bib ab 1-2	T t s3/7/1-2
<b>View Query</b>	D QUE d que L8	not available
<b>View Session History</b>	D HIS	DS
<b>Erase Entire Session</b>	DEL HIS	B

Adapted from handout prepared by David C. Brown, Dialog Account Representative

## Database Vendor Trivia

### *Dialog*

System is available 24 hours a day, 7 days a week, except from 1 AM to 7 AM Saturday.

Help Desk: 800-334-2564

24 hours M-F; 10 AM to 6 PM Saturday

Bluesheets for each database are available on the Web at:

<http://library.dialog.com/bluesheets>

Bluesheets are also available online in file 415.

### *STN*

System is available from noon on Sunday through Saturday at 6 PM.

Help Desk: 800-848-6533

8 AM to 8 PM, M-F

Database summary sheets for each database are available on the Web at:

<http://info.cas.org/ONLINE/DBSS/dbsslist.html>

FEBRUARY/MARCH 1997 NUMBER 54

## **BIOTECHNOLOGY & CHEMICAL NEWS**

### **ELEMENTS OF A GOOD SEARCH REQUEST**

\* *Describe in a few sentences what it is you are looking for.*

You may want to include keywords in these sentences and highlight them by underlining. Just listing keywords with no indication of how they relate to each other is often confusing, and may actually prevent the searcher from retrieving useful references. If separate answer sets are requested for broad terms, e.g. AIDS and AZT, without indicating any relationship between the terms, overwhelming and useless answer sets will be retrieved. A keyword might have a specific meaning to the examiner, but may actually be ambiguous if no frame of reference has been provided. Acronyms are particularly dangerous. Submitting lists of solitary keywords out of context also prevents the searcher from being able to select controlled vocabulary terms which describe the art and which could actually be more successful at retrieving relevant articles/patents than the keyword would be.

\* *Structure Searches.*

In the case of a structure search it is most cost effective to do a broad structure search first and then to limit to more specific or elected species in a subset search (cost=\$20) of the broad structure set. If a very specific structure search is done first and no answers are found, then broadening the search requires a completely new structure search costing another \$100.

\* *Claims and inventors.*

Many examiners include a copy of the abstract and/or claims with the request to improve the search process. If these items are well written, they can be very useful in helping the searcher to understand the claimed invention.

Including the inventors' names with the search request can improve the search in several ways. Inventors do sometimes forget to mention their own publications. An inventor search may be especially useful with foreign applicants when English language equivalents are sought. Many inventors publish the claimed invention after filing. Retrieving this reference helps the searcher gain insight into appropriate controlled vocabulary and indexing terms to use to search the subject of interest. The best way to provide the inventors' names is to attach a copy of the case cover to the search request form. This reduces the chance of spelling errors and also gives the searcher information on the filing and priority dates you may be trying to beat.

\* *Database Selection.*

There are over 700 databases available on-line. In almost all cases the searcher is the one who is best able to determine in which databases the claimed subject matter is likely to be found. However if you have always found good references in a particular database or if you know that a certain database indexes a particular journal which is pertinent to your art, it may be helpful to mention that database in your search request. These suggestions should be limited to one or two databases. Routinely requesting that searches be done in a large group of databases should be avoided; it results in cumbersome, costly and inefficient searches. On the other hand, strictly delineating the databases to be searched may be too limiting, eliminating databases which might contain useful information.

Again if you have any questions about how to write your search request or how to interpret results don't hesitate to talk with one of the searching staff. John Dantzman 308-4488, Jan Delaval 308-4498, Kathleen Fuller 308-4290, Mary Hale 308-4258, Barb O'Bryen 308-4291, David Schreiber 308-4292, Beverly Shears 308-4994, Paula Sheppard 308-4499, Mark Spencer 308-4266, Alex Waclawiw 308-4261.

Kathleen Fuller and Dilip Pandya prepared this newsletter.

